

Human heuristics for AI-generated language are flawed

Year: 2022 | Citations: 242 | Authors: Maurice Jakesch, Jeffrey T. Hancock, Mor Naaman

Abstract

Significance Human communication is now rife with language generated by AI. Every day, across the web, chat, email, and social media, AI systems produce billions of messages that could be perceived as created by humans. In this work, we analyze human judgments of self-presentations written by humans and generated by AI systems. We find that people cannot detect AI-generated self-presentations as their judgment is misguided by intuitive but flawed heuristics for AI-generated language. We demonstrate that AI systems can exploit these heuristics to produce text perceived as “more human than human.” Our results raise the question of how humanity will adapt to AI-generated text, illustrating the need to reorient the development of AI language systems to ensure that they support rather than undermine human cognition.